

## Ferramenta computacional de auxílio à interpretação de exames laboratoriais

André N. Medeiros<sup>1†</sup>, Ricardo M. Salgado<sup>2</sup>

<sup>1</sup> Discente do curso de Ciência da Computação pela Universidade Federal de Alfenas.

<sup>2</sup> Docente do Instituto de Ciências Exatas, Universidade Federal de Alfenas.

**Resumo:** A interpretação de exames laboratoriais bioquímicos frequentemente apresenta desafios significativos devido à diversidade de formatos e estruturas adotadas por diferentes laboratórios. A variação na apresentação dos dados e a falta de padronização muitas vezes dificultam a análise eficiente desses resultados, podendo levar a erros de interpretação e diagnósticos imprecisos. Profissionais da saúde enfrentam um obstáculo ao lidar com essa heterogeneidade, o que pode afetar a qualidade e a rapidez no processo de tomada de decisões clínicas. Nesse contexto, esse trabalho se concentra no desenvolvimento de uma ferramenta inovadora para lidar com essa problemática. Uma aplicação foi concebida utilizando tecnologias avançadas como Reconhecimento Óptico de Caracteres (OCR) combinado com uma Rede Neural Recorrente do tipo Long Short-Term Memory (LSTM). Essa abordagem permite a leitura precisa de exames provenientes de diferentes laboratórios, independentemente dos formatos utilizados. A aplicação se destaca ao criar uma interface que padroniza os resultados obtidos de maneira a simplificar a interpretação dos exames, facilitando assim o processo diagnóstico por esses profissionais. O modelo desenvolvido representa uma solução eficaz para os problemas expostos, superando a complexidade decorrente da diversidade de formatos. Ao automatizar e simplificar a análise, ela contribui para a redução de erros de interpretação e, conseqüentemente, para uma melhoria na qualidade do diagnóstico, impactando positivamente a eficiência e a precisão na área de saúde.

**Palavras-chave:** Análise laboratorial; Ferramenta; Redes Neurais; Tesseract; OCR.

## Computational tool to aid in the interpretation of laboratory tests

**Abstract:** The interpretation of biochemical laboratory tests often presents significant challenges due to the diversity of formats and structures adopted by different laboratories. The variation in data presentation and lack of standardization often hinder the efficient analysis of these results, potentially leading to misinterpretation and imprecise diagnoses. Healthcare professionals face an obstacle when dealing with this heterogeneity, which can affect the quality and speed of the clinical decision-making process. In this context, this work focuses on developing an innovative tool to address this issue. An application was devised using advanced technologies such as Optical Character Recognition (OCR) combined with a Long Short-Term Memory (LSTM) Recurrent Neural Network. This approach enables precise reading of tests from different laboratories, regardless of the formats used. The application stands out by creating an interface that standardizes the results in a way that simplifies the interpretation of the tests, thereby facilitating the diagnostic process for these professionals. The developed model represents an effective solution to the exposed problems, overcoming the complexity arising from the diversity of formats. By automating and simplifying the analysis, it contributes to reducing interpretation errors and, consequently, improving the quality of diagnoses, positively impacting efficiency and precision in the healthcare field.

**Keywords:** Laboratory analysis; Tool; Neural Networks; Tesseract; OCR.

---

<sup>1†</sup> Autor correspondente: [andrenmed@gmail.com](mailto:andrenmed@gmail.com).

## Introdução

Desde o alvorecer da humanidade, nossa espécie tem buscado compreender e explorar o mundo ao nosso redor. Em oposição ao aceite passivo de explicações até então transcendentais, o desejo intrínseco de conhecimento culminou na criação da ciência, uma ferramenta que permitiu à humanidade investigar e desvendar os segredos da natureza. Ao longo dos séculos, a ciência evoluiu e se diversificou, nesse meio, a computação surgiu como uma aliada fundamental, impactando significativamente todas as áreas do conhecimento, contribuindo para nossa compreensão do mundo mais uma vez.

A revolução tecnológica impulsionou o desenvolvimento de soluções computacionais que transcendem as fronteiras das disciplinas tradicionais. A computação se tornou uma ferramenta onipresente e indispensável, oferecendo apoio em diversas áreas, da astronomia à economia, da engenharia à educação. Essas inovações tecnológicas também desempenharam um papel fundamental na medicina.

A incorporação de avanços tecnológicos na área da saúde permitiu a criação de ferramentas e sistemas inovadores, como a aplicação desenvolvida neste estudo, destinada a auxiliar na interpretação de exames laboratoriais. A introdução de algoritmos de processamento de imagens e técnicas de aprendizado de máquina possibilitou uma análise mais precisa e rápida, tornando viável o diagnóstico através de exames complexos em um curto período de tempo.

Além disso, a evolução da tecnologia proporcionou a digitalização e o armazenamento seguro de dados clínicos, garantindo o acesso a informações relevantes sobre os pacientes de forma imediata e precisa. A contínua inovação nessa área reforça o compromisso em fornecer soluções cada vez mais sofisticadas, melhorando a qualidade e a eficácia dos serviços de saúde com foco na melhoria da vida e do bem-estar dos pacientes.

A interpretação de exames laboratoriais, em especial os exames bioquímicos, é uma etapa crítica na prática clínica, fundamental para o diagnóstico e acompanhamento de pacientes. No entanto, esse processo muitas vezes se torna desafiador devido à complexidade, à quantidade de informações contidas nesses documentos e a falta de padronização do formato pelos laboratórios que os realizam. Profissionais de saúde, como médicos, enfermeiros e biomédicos, frequentemente enfrentam dificuldades na análise desses resultados, o que pode impactar diretamente a qualidade do atendimento prestado e a tomada de decisões clínicas.

Nesse contexto, este artigo se concentra no desenvolvimento de um modelo com o objetivo central de abordar uma lacuna significativa no campo da saúde. A pesquisa propõe o desenvolvimento de uma ferramenta inovadora voltada para auxiliar os profissionais de saúde na interpretação de exames laboratoriais, com foco mais específico em exames bioquímicos, como os exames de sangue. O intuito é atender à urgência de aprimorar a análise desses exames, tornando-a mais eficiente e principalmente acessível, visando assim a melhoria da prática clínica e no atendimento do paciente.

A abordagem adotada nesta investigação envolveu a concepção e a implementação de uma ferramenta que combina conhecimentos de computação, processamento de dados e bioquímica. Por meio dessa abordagem multidisciplinar, foi pretendida a criação de um instrumento que permita aos profissionais de saúde a obter leituras mais rápidas e compreensíveis dos exames, facilitando, assim, a interpretação dos resultados e a tomada de decisões clínicas embasadas em evidências sólidas.

O presente trabalho encontra-se organizado da seguinte forma: na seção “Revisão Bibliográfica” é apresentada a fundamentação de conceitos básicos para contextualização do estudo. Posteriormente são apresentados os métodos utilizados e suas respectivas etapas, descrevendo a base de dados, preparação, métricas utilizadas, os modelos propostos e suas características. Por fim, são descritos a discussão dos resultados obtidos, a conclusão do trabalho e sugestões para futuros estudos.

## Revisão Bibliográfica

A utilização de tecnologias como Redes Neurais e Reconhecimento Óptico de Caracteres (OCR) tem sido amplamente explorada em diversos campos, oferecendo soluções mais eficientes frente aos métodos

tradicionais. Essas tecnologias são aplicadas de forma abrangente, tanto em contextos gerais quanto na área da saúde, mostrando-se valiosas em diferentes domínios.

Uma das aplicações importantes que utilizam o OCR está no setor bancário, em que é utilizado no processamento de cheques sem a intervenção humana. Um cheque pode ser inserido em uma máquina, onde o sistema escaneia o valor a ser emitido e realiza a transferência no valor correto. “Essa tecnologia foi quase que levada à perfeição para cheques impressos, e possui precisão mais que suficiente para cheques escritos à mão, reduzindo os tempos de espera em bancos.” (SINGH; BACCHUWAR; BHASIN, 2012).

O CAPTCHA (*Completely Automated Public Turing test to tell Computers and Humans Apart*, ou Teste de Turing público completamente automatizado para distinguir entre computadores e pessoas) é outro programa que faz uso dessa tecnologia. Se trata de uma aplicação que gera testes que os humanos podem passar, mas os computadores atuais não. O *hacking* é uma séria ameaça ao uso da internet hoje em dia, atualmente a maior parte das atividades humanas como transações bancárias, admissões em instituições de ensino, registros no geral, reservas, etc., são realizadas pela internet, todas essas atividades requerem uma senha que é muitas vezes maliciosamente utilizada por invasores. Eles desenvolvem programas, como ataques de dicionário e inscrições automáticas falsas, o que resulta no desperdício de memória e recursos dos sites. O ataque do dicionário, por exemplo, é uma tentativa contra sistemas com autenticação de senha, onde um invasor escreve um programa para tentar repetidamente diferentes senhas, como as provenientes de um dicionário das senhas mais comuns. No CAPTCHA, uma imagem composta por uma série de letras ou números é gerada e é obscurecida por meio de técnicas de distorção de imagem, variação de tamanho e fonte, fundos distrativos, segmentos aleatórios, destaques e ruído na imagem. Esse sistema pode ser empregado para eliminar esse ruído e segmentar a imagem, tornando-a mais acessível para os sistemas de OCR.

Uma das principais atividades realizadas pelos profissionais da saúde é a tomada de decisões. Isso acontece em diversos pontos de sua rotina e são baseadas em uma série de informações coletadas em diferentes momentos, uma delas é a de diagnóstico. Uma grande contribuição da computação para o campo médico diz respeito ao diagnóstico, com o fornecimento de um conjunto de informações que os usuários precisam para realizar uma interpretação, isso de forma rápida e no momento que eles precisam.

A primeira experiência bem-sucedida de se construir um sistema de diagnóstico médico foi o programa MYCIN [Shortliffe, 1976], desenvolvido nos Estados Unidos no início da década de 1970 pelo Dr. Edward Shortliffe, da Universidade de Stanford. Esse sistema tem a função de apoio ao diagnóstico e terapia de infecções provocadas por microrganismos. (AMBROSIO, 2002, p. 53).

Essa área também tem testemunhado um aumento no uso da tecnologia OCR para processamento de documentos. Profissionais da saúde frequentemente lidam com grandes volumes de documentos para cada paciente. Acompanhar todas essas informações pode ser custoso em relação ao tempo, por isso é útil inserir os dados mais relevantes em um banco de dados eletrônico que possa ser acessado conforme necessário. Ferramentas de processamento de documentos, impulsionados pela tecnologia OCR, são capazes de extrair informações desses laudos e inseri-las em um banco de dados, garantindo que estes dados sejam prontamente registrados e mais facilmente acessíveis.

As aplicações dessas tecnologias são vastamente conhecidas pela sua versatilidade e inúmeras aplicações em vários campos, desde a gestão de documentos até o diagnóstico médico. A literatura científica reflete seu impacto positivo em diversos aspectos, tanto no contexto geral quanto no específico da medicina.

Entretanto, apesar do uso consolidado dessas ferramentas em várias áreas, uma lacuna significativa persiste na literatura acadêmica. A ausência de um modelo direcionado para o processamento de exames laboratoriais prontos, visando ser uma ferramenta de apoio clínico para profissionais de saúde, com utilização do OCR para essa finalidade, é um ponto que carece de atenção.

Até o momento, os estudos e aplicações exploradas frequentemente focam em outras áreas médicas ou em procedimentos específicos, deixando uma brecha considerável na adaptação direta dessas tecnologias para a interpretação e análise de exames laboratoriais. Portanto, existe um espaço substancial para o desenvolvimento e aprimoramento de um modelo direcionado a esse fim, o que poderia impactar positivamente

a prática clínica, oferecendo suporte aos profissionais de saúde e agilizando a interpretação desses documentos vitais.

Essa lacuna identificada na literatura acadêmica ressalta a necessidade premente de uma abordagem direcionada para a aplicação do OCR e Redes Neurais na interpretação de exames laboratoriais, uma ferramenta capaz de otimizar a rotina clínica e melhorar a eficiência no diagnóstico e tratamento dos pacientes.

## Fundamentação teórica

De acordo com Andriolo et al. (2018), o diagnóstico laboratorial é um campo em constante crescimento, desempenhando um papel de extrema importância na orientação da tomada de decisões clínicas. Ele oferece suporte essencial não apenas para a prevenção, mas também para o diagnóstico e monitoramento terapêutico de uma ampla gama de patologias humanas, abrangendo virtualmente todas as áreas da medicina. Na literatura, autores como Lima (2017) irão dividir esse processo todo em 3 grandes etapas: pré-analítica, que é a coleta, transporte e processamento das amostras; analítica, que é a realização do teste em si; e pós-analítica, que consiste na leitura e interpretação dos resultados, justamente a lacuna em que este trabalho busca preencher.

Um dos problemas relacionados ao atraso ou prolongamento dessa terceira fase é o tempo de leitura prévio por parte do profissional, para interpretar e diagnosticar possíveis problemas no ato da consulta. A computação se alinha nesse momento para catalisar esse processo de diagnóstico. Por meio de tecnologias de OCR é possível obter as informações chave de um exame de forma mais rápida e estruturada. Esse tipo de tecnologia, embora mais antiga, se alinha nos dias de hoje à redes neurais artificiais, o que aumenta a capacidade desse tipo de ferramenta.

Para Nagy (1991, p.1, tradução nossa) “O reconhecimento óptico de caracteres debutou comercialmente em 1955 devido a melhorias nos processadores da época e ao desenvolvimento de tecnologias de digitalização”. Os avanços seguintes vieram como melhorias dos métodos usados para “reconhecer” os documentos, junto com a ideia de usar o contexto do conteúdo escrito para melhorar esse reconhecimento. Esse novo sistema automatizou o processo de passar documentos escritos ou impressos novamente para caracteres reconhecidos. O autor, sendo um membro Sênior da IEEE (Instituto de Engenheiros Elétricos e Eletrônicos, responsável pela criação de diversas normas e padrões) escreveu um livro chamado “*At the Frontiers of OCR*” que falava sobre o limite da tecnologia e suas implicações para a época vai dizer: “Um único sistema costumava realizar o trabalho de dez a vinte operadores de entrada de dados” (NAGY, 1992, p. 1, tradução nossa).

Ademais, um outro detalhe é que a venda desses sistemas era feita de forma personalizada, um conjunto específico de formatos e de fontes deveriam ser especificadas antes que pudesse ser usado. As limitações computacionais da época tornavam o processo de reconhecimento de textos muito custoso em relação ao tempo, pois regras específicas deveriam ser manualmente inseridas dependendo do caso.

Cada caractere era processado sequencialmente e independentemente. Somente o pré-processamento (escala de cinza, normalização, eliminação de ruído, identificação de linhas e definição de alpha/numéricos) e pós-processamento (verificação ou correção ortográfica, às vezes incorporando regras léxicas personalizadas) eram grandes entidades levadas em consideração. O legado desse período, o paradigma da Localização de Caractere e Isolamento de Caractere ainda anda conosco. (NAGY, 1992, p. 1, tradução nossa).

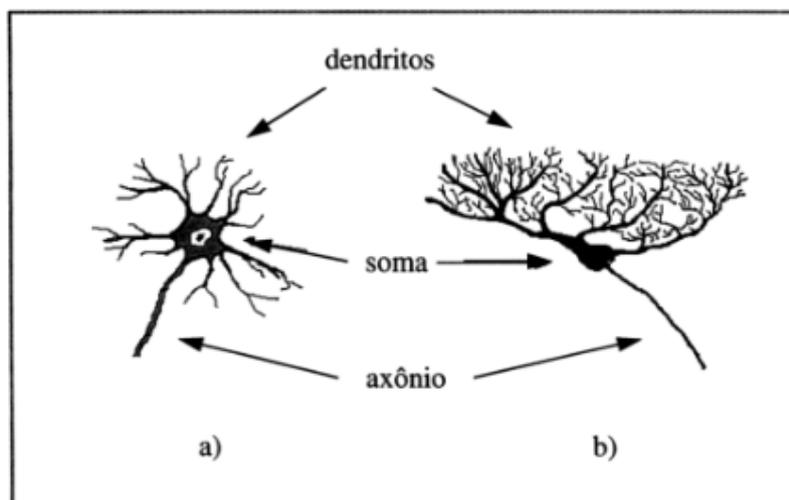
Esses problemas foram sanados com os avanços constantes da área de Inteligência Artificial, especialmente em Redes Neurais Artificiais (RNAs). As RNAs estão fortemente atreladas às principais ferramentas para reconhecimento de caracteres atualmente, por isso mostra-se necessário a compreensão básica de uma RNA.

## Redes Neurais

Todo órgão do corpo humano é composto por células, mas as células que compõem o sistema nervoso são especiais, e é pelo seu funcionamento extremamente complexo que surge a percepção que se tem do mundo, as emoções e a consciência humana no geral.

A célula nervosa, ou neurônio como é muitas vezes chamado, foi identificado pelo neurologista espanhol Ramón y Cajal no século XIX. Um neurônio é composto pelo corpo celular, conhecido como “soma”, o Axônio e os Dendritos. O axônio é o prolongamento do corpo celular que é responsável por transferir informações pelo sistema nervoso, eles podem variar de milímetros até um metro de comprimento. No final de um axônio fica o terminal axonal, que é onde ele se conecta com os dendritos de outro neurônio para que a transmissão possa ocorrer. Os dendritos também são prolongamentos do corpo celular, mas eles servem somente para receber sinais. A Figura 1 ilustra um neurônio com suas partes apontadas.

Figura 1. Neurônios do sistema nervoso central do vertebrados: a) neurônio motor da medula espinhal; b) célula de Purkinje encontrada no cérebro. Esta célula é notável pela extensa ramificação da sua árvore dendrital da qual apenas uma pequena parte é mostrada na figura, podendo ultrapassar várias dezenas de vezes as dimensões do corpo celular.



Fonte: Adaptado de Kovács (2006).

Sinapse é como é chamado o ponto de contato entre dois neurônios, é composta de um lado pré sináptico (onde é enviado o sinal) e um lado pós sináptico (onde é recebido), a chamada transmissão sináptica é a transferência da informação de um neurônio para outro. Quando um impulso elétrico chega ao final do axônio são liberadas substâncias químicas chamadas de neurotransmissores, eles são detectados pelo dendrito que se encontra do outro lado da sinapse. Ao ser detectado, é gerado no dendrito um impulso elétrico que faz com que o processo se repita.

A grosso modo, podemos dizer que as sinapses são regiões eletroquimicamente ativas, compreendidas entre duas membranas celulares: a membrana pré-sináptica, por onde chega um estímulo proveniente de uma outra célula, e a membrana pós-sináptica, que é a do dendrito. Nesta região inter-sináptica, o estímulo nervoso que chega a sinapse é transferido à membrana dendrital através de substâncias conhecidas como neurotransmissores. O resultado dessa transferência é uma alteração no potencial elétrico da membrana pós-sináptica. Dependendo do tipo de neurotransmissor, a conexão será excitatória ou inibitória<sup>2</sup>. (KOVÁCS, 2006, p.15)

<sup>2</sup> Conexão excitatória é a que contribui para formação de um impulso nervoso no axônio de saída, a outra o oposto.

A maioria dos axônios é envolvida por uma camada de proteínas chamada de Bainha de Mielina, ela isola grande parte do axônio fazendo com que seja acelerada a transmissão do impulso elétrico pela membrana do neurônio. A Esclerose Múltipla por exemplo, faz com que o sistema imune ataque essa estrutura, o que interrompe a comunicação entre vários neurônios e diversas lesões surgem no sistema nervoso, impactando muito – e de diversas formas – as capacidades cognitivas da pessoa.

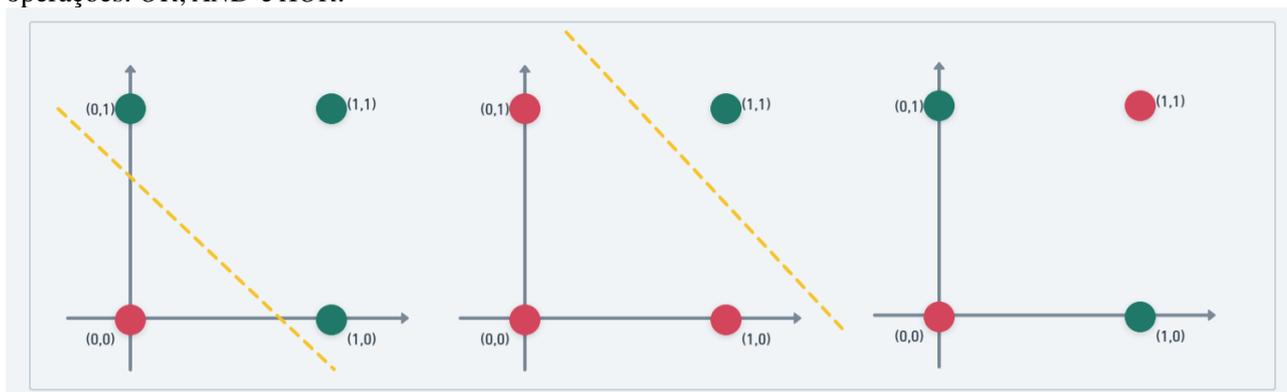
Segundo Kovács (2006), o primeiro estudo publicado no campo de redes neurais artificiais envolveu a criação do chamado Neurônio Booleano de McCulloch, criado sob inspiração dos resultados de Alan Turing e John von Neumann. Eles acreditavam que a inteligência era booleana<sup>3</sup>, Warren McCulloch publica no Boletim de Biofísica Matemática um artigo com o título (em tradução livre) “Um Cálculo Lógico de Ideias Inerentes à Atividade Nervosa”, que é a base até os dias de hoje para as teorias de RNAs. Apesar de toda a diferença temporal e limitações da época, o neurônio de McCulloch era extremamente simples: um dispositivo binário em que sua saída pode ser “pulso” ou “não pulso”, e as entradas possuíam um ganho arbitrário, podendo ser excitatórias ou inibitórias. A determinação da saída era dada pela soma ponderada das entradas com os respectivos ganhos como fatores de ponderação (positivos quando excitatórios e negativos quando inibitórios), caso esse resultado fosse maior ou igual a um limiar definido, a saída era “pulso”, caso contrário, “não-pulso”.

Esse estudo foi importante para a concepção do neurônio Perceptron entre 1957 e 1958, como também para sucessivos modelos como o Adaline em 1960. O perceptron na época não passou de uma proposta, pois não existiam recursos que permitissem que ele fosse ajustado e nem um algoritmo que pudesse sustentá-lo. Foi uma proposta que resolvia problemas extremamente simples. O Perceptron chegou a ser implementado na máquina Mark I pela primeira vez e foi o pioneiro em reconhecimento de imagens, mas por todos os motivos mencionados foram deixados de lado temporariamente.

O exemplo clássico que demonstra as limitações do Perceptron é o problema do “ou exclusivo” (*XOR*), isso porque ele só pode aprender a separar classes linearmente. O “ou exclusivo” é um operador lógico que retorna verdadeiro somente se os operandos forem diferentes (um verdadeiro e um falso). Se houver um conjunto de dados que represente o *XOR* como  $(0,1)$ ,  $(1,0)$ , então não é possível traçar uma única linha para separar as duas classes distintas  $(0$  e  $1)$ .

Na figura 2 abaixo se observa o comportamento de cada operação, respectivamente: “Ou”, “E” e “ou exclusivo”.

Figura 2: Demonstração visual do problema do *XOR*, da esquerda para a direita vê-se respectivamente as operações: *OR*, *AND* e *XOR*.



Fonte: Dos autores.

Em um espaço bidimensional, o perceptron pode criar um hiperplano para separar duas classes que podem ser divididas por uma única linha. Contudo, no caso do *XOR*, as classes não são linearmente separáveis;

<sup>3</sup> Tipo de dado primitivo no campo da Ciência da Computação, possui dois valores: 0 (falso) e 1 (verdadeiro). O nome é em homenagem a George Boole que definiu o sistema de lógica algébrica pela primeira vez no século XIX.

portanto, o perceptron não consegue encontrar uma solução para esse problema. Em outras palavras, não é possível com uma única linha segmentar as classes 1 e 0.

A crítica sobre o Perceptron foi tão grande que o estudo de redes neurais foi deixado de lado dando contribuindo para o início do chamado “Inverno da IA”, período marcado pela falta de desenvolvimento na área de inteligência artificial por baixo financiamento, limitação teóricas no funcionamento de redes neurais e inexistência de hardware adequado para seu suporte.

A solução para o problema XOR já era conhecida. Bastava acrescentar mais uma camada de neurônios na rede (uma camada escondida). O que faltava era um algoritmo que fosse capaz de treinar os pesos dessa rede multi-camada para que pudesse classificar corretamente problemas mais complexos. Várias soluções equivalentes foram descobertas durante os anos seguintes, mas só a publicação do algoritmo de “retropropagação de erro” (error backpropagation) por [Rumelhart et al., 1986] popularizou uma solução de carácter universal para esse tipo de problema. A partir desse momento, surgiram os modelos que foram desenvolvidos durante os anos tranquilos da pesquisa e inúmeros outros modelos de redes neurais artificiais junto com algoritmos de aprendizagem foram apresentados. (RAUBER, 2005, p.4)

Atualmente, as redes neurais artificiais representam uma das áreas mais promissoras e inovadoras no campo da inteligência artificial. Com avanços significativos no poder computacional, conjuntamente com técnicas de aprendizado profundo, as redes neurais conquistaram aplicações amplas e impactantes em diversas esferas. Desde reconhecimento de fala e visão computacional até diagnósticos médicos assistidos por computador, as redes neurais tornaram-se a espinha dorsal de muitos avanços tecnológicos.

## OCR

Um módulo de leitura de PDF poderia ter sido escolhido para obtenção dos dados no desenvolvimento do presente trabalho, mas isso iria restringir o funcionamento, dado que era um objetivo inicial que imagens (fotos dos exames) pudessem também ser submetidas. O OCR é uma tecnologia que permite que textos sejam extraídos de documentos no geral. Ela depende de alguns fatores para obter uma melhor leitura, os principais são o idioma, a qualidade da entrada e tipo de fonte utilizada.

No mercado, esse tipo de tecnologia é utilizada para automatização de processos que a princípio precisavam de entradas manuais. De forma análoga, foi usado aqui como processamento inicial da entrada.

O OCR pode ser pensado como uma compressão, basicamente pega-se uma imagem em duas dimensões e com alguma profundidade, então essa imagem é passada por vários estágios de compressão para extrair somente a informação necessária, no caso, o texto.

O histograma<sup>4</sup> de uma imagem normal possui frequências bem distribuídas (figura 3), o primeiro passo do OCR é binarizar a imagem. A binarização é o método mais simples de segmentação de imagens, basicamente consiste em separar uma imagem em regiões de interesse e não interesse através de um ponto de corte. A definição desse ponto de corte é chamado de limiarização, que define um único ponto chamado de *threshold*. O resultado numa imagem comum é um único componente conectado, ou em outras palavras, uma única borda, separando o objeto do fundo, isolando o conteúdo desejado, a figura 4 mostra uma imagem de entrada sem ser processada, a figura 5 mostra essa mesma imagem, após passar pelo processo de binarização. O resultado desse processo “limpa” a imagem, tornando-a livre de quaisquer ruídos que possam vir a atrapalhar a leitura posterior do OCR.

Trazendo para o contexto deste trabalho, após as etapas descritas a aplicação resultaria em uma imagem de um caractere de forma isolada, ou uma sequência de caracteres isolados, mas mesmo após esse estágio, a identificação da letra específica do alfabeto ainda não estaria determinada. O próximo passo é a Classificação. Conceitualmente, o OCR faz uma classificação do alfabeto por correspondência de padrões, mas isso varia muito de acordo com a fonte da imagem, por exemplo. É na etapa de classificação que conceitos de Redes Neurais ou Deep Learning começam a ser aplicados em OCRs específicos, como é o caso do

<sup>4</sup> Espécie de gráfico de barras que demonstra uma distribuição de frequências de 0 a 255, do preto ao branco.

Tesseract. Sucintamente, a classificação é uma comparação da imagem de um caractere com sua versão conhecida.

Resumindo, partindo de uma imagem binarizada se formam componentes conectados e cada um deles irá se assimilar com caracteres do alfabeto, para termos a informação, é preciso classificá-los. Para isso, é comparado cada componente com caracteres conhecidos, seja por correspondência ou por um conjunto de treino (dados apresentados ao modelo para criação e ajuste do modelo), e com isso é tido em certo nível, uma medida, um valor numérico que indica o quão correspondente uma coisa é da outra. Mas não é só isso, já que na maioria das vezes não se tem um caractere isolado, mas sim uma sequência (que forma uma palavra). Nesse caso, não é aceito imediatamente o maior valor de correspondência. Supondo um vetor de correspondência de um componente segmentado (um caractere) com diferentes fontes.

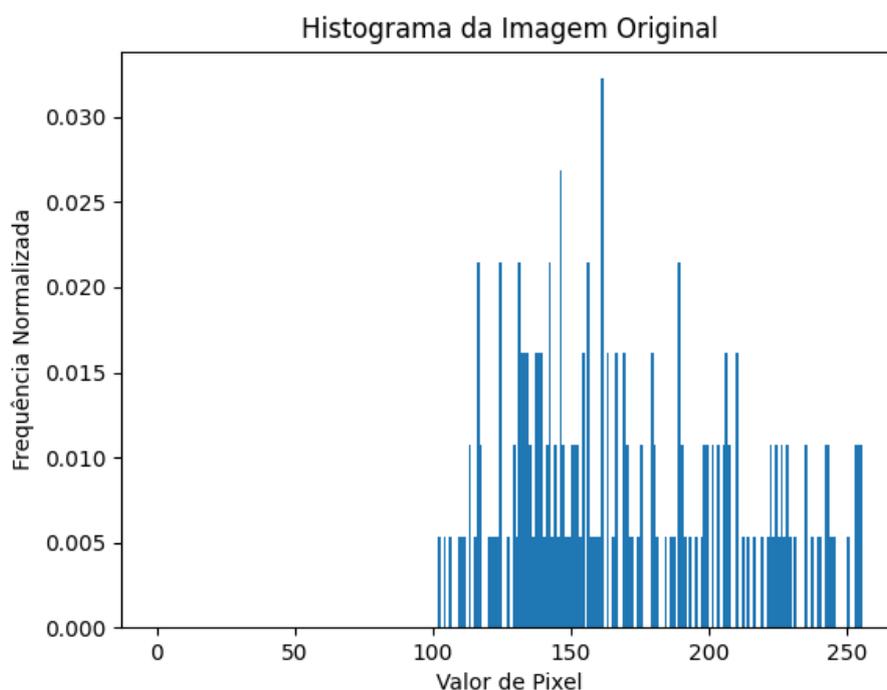
$$V = [0.2, 0.7, 0.89, 0.98]$$

A medida de correspondência do caractere com a quarta fonte (V[3]) foi a maior, mas existindo outros componentes, será feita a verificação com todos os caracteres dessa palavra. Sendo uma palavra de 3 letras, se tem uma matriz, como:

$$M = [[0.2, 0.7, 0.89, 0.98], [0.1, 0.7, 0.95, 0.7], [0.2, 0.5, 0.89, 0.65]]$$

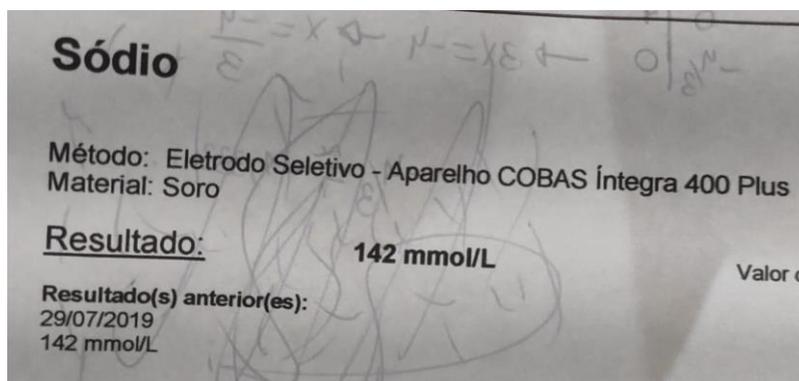
Então é usado estatística populacional para definir qual é a fonte que de fato se aplica e dessa forma melhorar o processamento. Após calcular a média para cada fonte (itens da mesma posição em cada linha), se têm aproximadamente os valores: 0.166, 0.633, 0.91, 0.776. Portanto, o conjunto de caracteres conhecido que mais se encaixa com a leitura feita é o terceiro (V[2]), indicando que determinada fonte tem melhor correspondência. Especificamente no caso do Tesseract, é usada uma Rede Neural Artificial chamada LSTM.

Figura 3. Histograma de uma imagem comum.



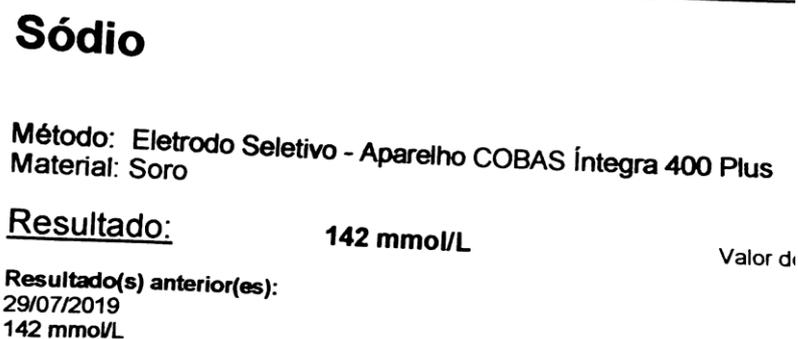
Fonte: Dos autores.

Figura 4. Imagem de entrada do processo de binarização.



Fonte: Dos autores.

Figura 5: Imagem de saída do processo de binarização.



Fonte: Dos autores.

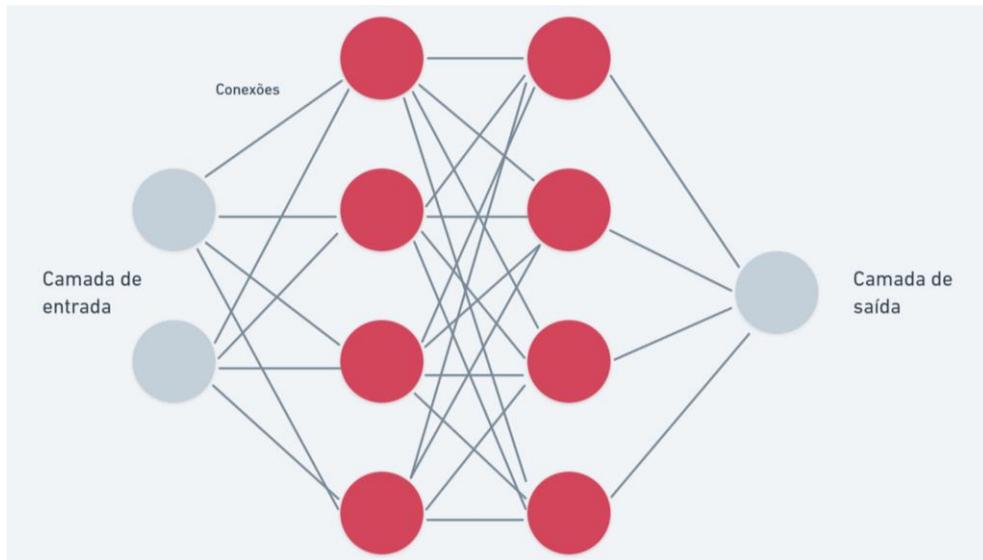
## LSTM

Para melhor entendermos o que é LSTM, primeiro precisamos saber o que é uma Rede Neural Recorrente, ou RNN. As RNNs são redes neurais utilizadas principalmente em processamento de linguagem natural (vertente da IA que ajuda os computadores a entender, interpretar e manipular a linguagem humana) em ferramentas como o autocomplete do Google, ou em assistentes virtuais.

Redes neurais artificiais simples não são capazes de realizar bem algumas tarefas. Vamos supor um modelo de tradução de línguas, em que a entrada é uma frase em inglês e a saída em português. Uma rede neural simples não é capaz de realizar essa tarefa pois possui um número fixo de neurônios de entrada e de saída, de forma que não podemos alimentá-lo com dados de entrada de tamanhos variáveis já que o modelo só aceita entradas de comprimento fixo. Além disso, o arranjo das palavras da frase importa, ou seja, a ordem em que as palavras aparecem na frase podem mudar seu significado durante a tradução, o que não é considerado por uma rede neural comum (sequencial), que independentemente da ordem em que as entradas ocorrem, irá reproduzir sempre um mesmo resultado. Por esses dois motivos, existe uma limitação desse tipo de modelo para determinados problemas, como é o caso do processamento de linguagem natural.

Resumindo, damos a uma rede neural sequencial as diferentes *features* na entrada de uma única vez, que então são passadas para a(s) camada(s) oculta(s) e é gerado um resultado na camada de saída. A Figura 6 abaixo ilustra uma rede neural sequencial.

Figura 6 : Rede Neural Artificial (RNA sequencial).



Fonte: Dos autores.

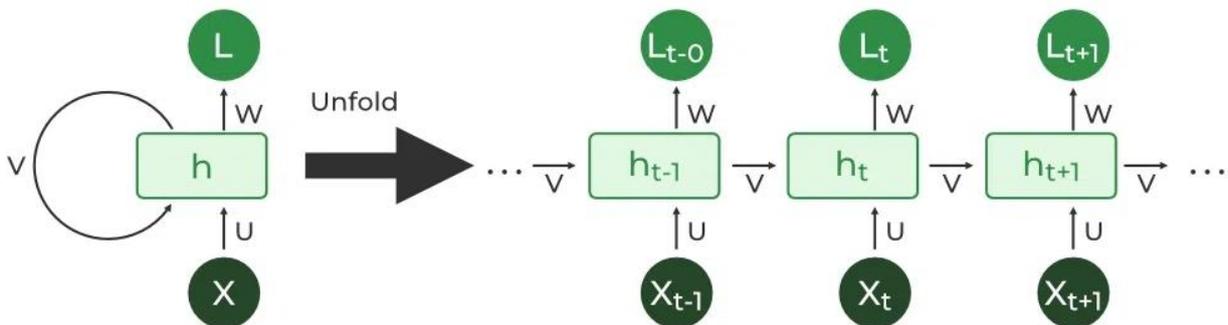
A rede neural recorrente tem esse nome pois ocorre repetidas vezes ao longo do tempo. Observando a primeira parte da figura 7 observamos a rede neural “enrolada”, na parte da direita vemos ela “desenrolada” indicando que a mesma rede neural é usada em função do tempo, cada seção desta é a mesma rede neural, entretanto as features são passadas uma a uma. Para cada passo, uma saída (output) é gerada e também é passada para a segunda instância e a segunda faz o mesmo por diante. Portanto, para todos os passos, exceto o primeiro, teremos duas entradas, uma delas sendo a saída da iteração anterior e a outra é a entrada desse passo específico.

O resultado de uma RNN é calculado matematicamente da seguinte forma:

$$y_t = \alpha(W_x X_t + W_y y_{t-1} + b)$$

Temos um conjunto de pesos da entrada representados por “W” e um outro conjunto de pesos para as saídas anteriores, por fim soma-se o *bias* (viés, peso adicional usado para ajustar a saída).

Figura 7: Rede Neural Recorrente (RNN) enrolada e desenrolada.



Fonte: Adaptado de GeeksForGeeks (2023).

No caso da LSTM ou em tradução livre, “Longa memória de curto prazo” existem para aplicações em que a entrada possui uma sequência ordenada, ou que o começo da entrada possa ser importante para processar o final. Ela é um tipo de RNN, portanto também usa a saída de uma rede em uma nova entrada. Nas LSTMs,

além de os nós serem importantes, eles também têm um estado interno. Cada nó usa o estado interno como um espaço de memória, e por isso a informação pode ser armazenada ou recuperada através de vários passos. O valor de entrada, a saída anterior e o estado interno são todos usados para calcular os nós, a nova saída gerada não serve só como entrada da próxima iteração, mas também para atualizar a memória interna.

Como toda RNA (rede neural artificial), os nós possuem parâmetros que determinam como as entradas são usadas nos cálculos, mas além desses as LSTMs têm parâmetros chamados “*gates*”, que controlam o fluxo da informação dentro do nó, particularmente o quanto a informação armazenada pode ou não ser “esquecida” pelos nós seguintes. Esses *gates* são pesos e *biases*, o que significa que o comportamento depende das entradas. Por exemplo, a entrada de uma letra “Q” não precisa de muita informação pois a próxima letra é quase certamente um “U”, mas para uma entrada “E” podemos precisar recuperar informações já passadas para que possa ser determinada a próxima. De forma análoga, existem *gates* que ditam o quanto a saída vai ser determinada pelos cálculos do nó e quanto vai ser determinada pela informação já salva. Apesar de parecer complexo, esses parâmetros são todos aprendidos durante o treinamento, assim como qualquer outra rede neural artificial.

Contudo, os desafios persistem, incluindo questões éticas, interpretabilidade dos modelos e a necessidade de lidar com conjuntos de dados desbalanceados. A busca por redes mais eficientes, robustas e éticas (garantindo que as decisões geradas pela rede neural sejam alinhadas com princípios como justiça, transparência, privacidade e equidade) permanece como um objetivo crucial para garantir o desenvolvimento responsável e o avanço contínuo desse campo. O futuro das RNAs está intrinsecamente ligado à colaboração interdisciplinar, que promove o entendimento mais aprofundado de como esses modelos funcionam e como podem ser aplicados de forma ética e eficaz, desempenhando um papel fundamental na resolução de problemas complexos e na inovação tecnológica.

## Desenvolvimento

### Visão geral

O desenvolvimento deste trabalho teve foco principalmente na criação da ferramenta de leitura dos exames bioquímicos. Após um levantamento detalhado das necessidades dos usuários e as características críticas, foram anotados os principais requisitos que o sistema deveria ter, que são estes:

- **Processamento automático:** a aplicação deveria ser capaz de identificar automaticamente os grupos de exames e seus valores a partir dos documentos enviados. Além disso, ela deveria também aplicar regras específicas de acordo com o laboratório identificado. Esse foi um ponto crítico do trabalho pois envolveu um esforço para selecionar os principais laboratórios e dessa forma ser capaz de atender a maior parte dos exames potencialmente processados.
- **Capacidade de Personalização:** a aplicação deveria ser capaz de permitir que os profissionais de saúde personalizem seus valores de referência individualmente para cada exame, uma vez que critérios institucionais ou de especialidade podem variar.
- **Visualização clara dos dados:** o usuário deve poder visualizar os dados graficamente, juntamente com valores máximos, mínimos e intervalo de referência.
- **Interface de usuário intuitiva:** pensando no uso amplo de uma aplicação desse tipo, a interface de usuário deve ser simples e intuitiva, para que os profissionais possam navegar facilmente pela aplicação, fazer upload de exames e interpretar os dados através de uma melhor visualização se comparada com o exame recebido inicialmente.

A solução criada é uma aplicação web full-stack que utiliza principalmente tecnologias oriundas do JavaScript, tanto para o front-end (cliente) quanto para o back-end (servidor), foram a biblioteca React.js (18.2.0) para criação da interface de usuário e o Express.js (4.18.2), framework do Node.js, que funciona como um conjunto mínimo e flexível da tecnologia mencionada. Além disso, para o banco de dados foi escolhido o MySQL, e o coração da aplicação é um motor de OCR (Optical Character Recognition, ou “Reconhecimento óptico de caracteres” em tradução livre) chamado Tesseract, implementado usando principalmente a linguagem Python (módulo PyTesseract).

Para garantir a integridade e o gerenciamento eficaz da aplicação durante o desenvolvimento, foi utilizado como sistema de controle de versão o Git (2005), juntamente com a plataforma GitHub. Permitindo o acompanhamento de mudanças e gerenciando o código de maneira organizada. O GitHub por sua vez, facilitou o compartilhamento do código e a integração contínua, garantindo que cada atualização fosse imediatamente incorporada ao projeto.<sup>5</sup>

## Base de Dados

Nesta pesquisa, os dados utilizados foram extraídos principalmente de documentos PDF, provenientes de diferentes laboratórios clínicos. Esses documentos continham uma variedade de exames bioquímicos comumente requisitados por profissionais de saúde para diagnóstico e acompanhamento de pacientes.

Foram selecionados documentos de laboratórios reconhecidos, como o laboratório da Unimed (Unilab, figura 8) e outros centros de análises de renome. Os exames escolhidos para a coleta de dados foram estrategicamente diversificados, abrangendo amostras de análises sanguíneas que abordam grupos corporais distintos.

Os documentos obtidos de diferentes laboratórios exibiam variações estruturais significativas. Cada instituição apresentava um formato único em seus relatórios, resultando em desafios para a padronização e extração uniforme dos dados para análise. Essas variações incluíam formatos distintos de apresentação de informações, disposição e até mesmo unidades de medida específicas adotadas pelos laboratórios.

Figura 8: Exemplo de exame analisado, laboratório Unilab.



**UniLab** Laboratório de Análises Clínicas

Av: Bento de Abreu, 1008 - Fonte - Araraquara/SP  
 Telefone: (16) 3303-1800 Unilab: 0800-113244  
 CNPJ: 45.272.366/0016-34 CNES: 6239447



---

**03-459766**

Nome Social: [REDACTED]

Nome: [REDACTED]

Origem: **UNILAB** Convênio: **UNIMED**

Médico Dr(a): [REDACTED]

Data/Hora Cadastro: 11/01/2023 08:03

Sexo/Idade: **M/20**

Data Nascimento: **25/04/2002**

---

### Hemograma Completo

Data/Hora da Coleta: **11/01/2023 - 08:03**

Método: Sistema Automatizado  
Material: Sangue

**Eritrograma**

<p>Eritrócitos: <b>5,41 milhões/mm<sup>3</sup></b></p> <p>Hemoglobina: <b>15,8 g/dl</b></p> <p>Hematócrito: <b>46,8 %</b></p> <p>V.C.M: <b>86,5 fl</b></p> <p>H.C.M: <b>29,2 pg</b></p> <p>C.H.C.M: <b>33,8 %</b></p> <p>R.D.W: <b>12,0 %</b></p> <p>Observação: Sem alterações morfológicas.</p>	<p>Valores de Referência:</p> <p>De 4,5 a 6,0 milhões/mm<sup>3</sup></p> <p>De 13,00 a 17,0 g/dl</p> <p>De 38 a 52 %</p> <p>De 82 a 98 fl</p> <p>De 27 a 32 pg</p> <p>De 32,0 a 36,0 %</p> <p>11,0 a 14,5 %</p>
---	---

Fonte: Dos autores.

Para o desenvolvimento do modelo, foi necessário o uso de um sistema de gerenciamento de dados relacional (MySQL) para armazenar e recuperar informações relevantes aos exames laboratoriais. O banco de

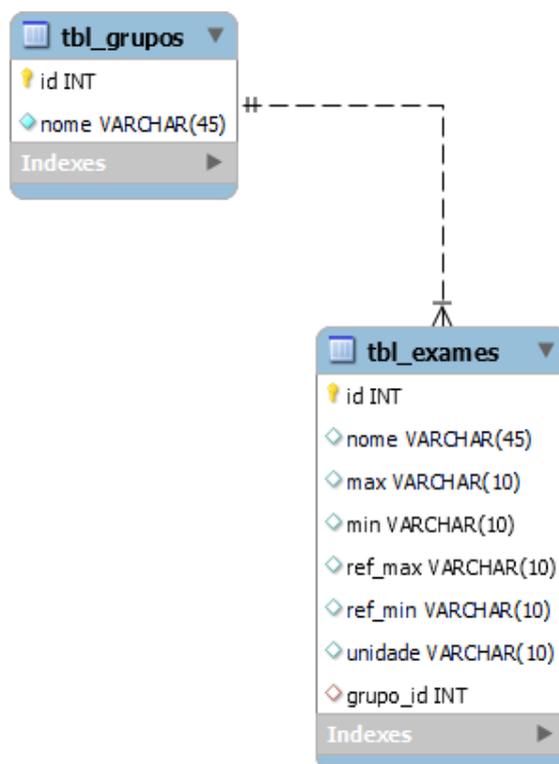
<sup>5</sup> O acesso ao projeto está disponível em: [github.com/SprigganCS/TCC](https://github.com/SprigganCS/TCC).

dados foi estruturado de forma que permitia a catalogação e a posterior interpretação dos dados obtidos dos documentos.

O esquema adotado para o banco de dados contemplou a criação de duas tabelas, uma que contém os grupos corporais e outra com os exames específicos dentro de cada categoria. A relação dos exames com o grupo correspondente foi feita através de uma chave estrangeira (muitos para um). A tabela de grupos, mais simples, possui um campo nome e um id, utilizado para relacionar os exames. A outra, além de um identificador para cada exame, possui colunas de valores máximos e mínimos (extrapolados), seus valores de referência e sua unidade.

A modelagem (figura 9) foi concebida para proporcionar flexibilidade e facilidade na inserção e recuperação dos dados, para adicionar um exame novo por exemplo, basta inserir na tabela de exames com os valores necessários e o sistema passará a procurar resultados desse exame mostrando na *dashboard* caso esteja disponível.

Figura 9: Modelo Entidade Relacionamento (EER) do esquema do banco de dados desenvolvido feito por engenharia reversa do MySQL Workbench.



Fonte: Dos autores.

## Interface e uso

A interface do sistema, desenvolvida com a biblioteca mantida pela Meta (2004), apresenta como principal característica a abordagem baseada em componentes. A arquitetura de componentes facilita significativamente a estruturação, manutenção e escalabilidade da aplicação, permitindo que os diferentes elementos da interface sejam encapsulados e reutilizados de maneira eficiente.

Além do React, o projeto contou com a utilização de várias bibliotecas e pacotes que enriqueceram a experiência do usuário e otimizaram o desenvolvimento. O Bootstrap foi adotado para estilizar a aplicação, ele fornece um conjunto de estilos e componentes pré-construídos que asseguram uma aparência consistente e profissional. O Axios é uma biblioteca que foi empregada para realizar as requisições HTTP de forma

simplificada e eficiente, garantindo uma comunicação eficaz entre o front-end e o back-end. Para gerenciar as rotas e a navegação na aplicação, o pacote React Router DOM foi incorporado, permitindo a criação de rotas e a navegação entre diferentes partes da aplicação de forma suave e responsiva. Por fim, foram utilizados o React Icons e React Spinners para aprimorar a experiência do usuário, fornecendo ícones e elementos visuais interativos que contribuíram para a usabilidade da aplicação.

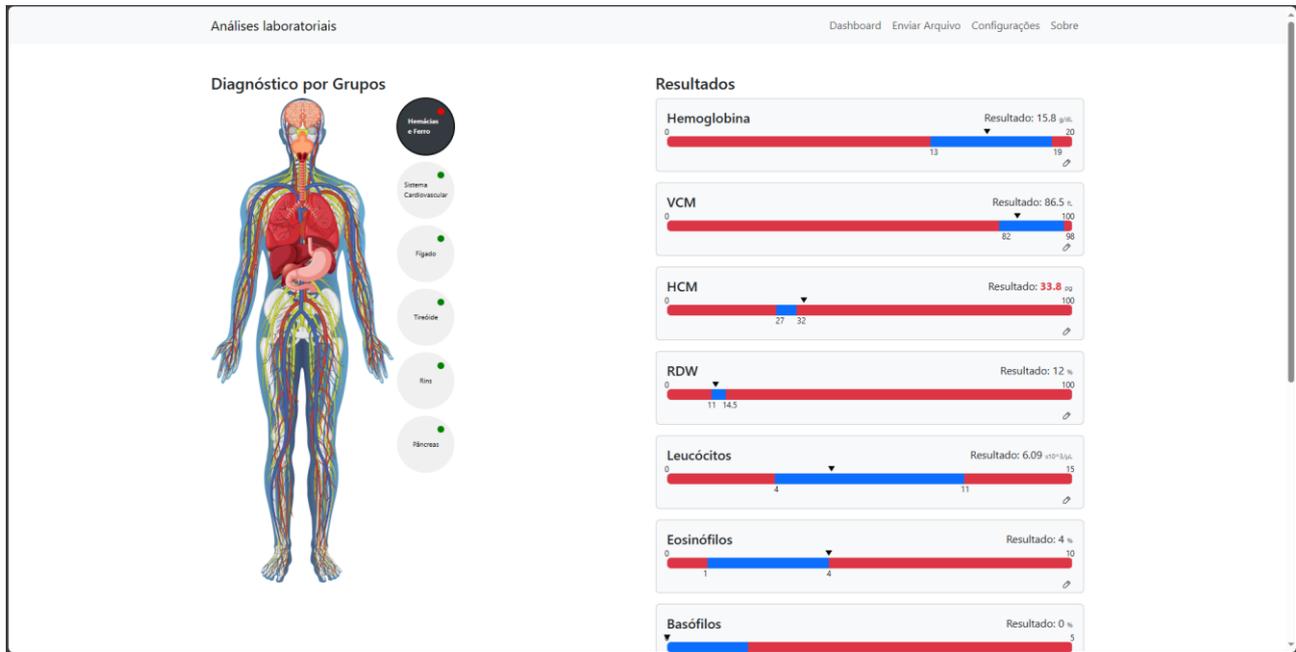
Todas essas tecnologias e pacotes desempenham um papel muito importante na construção da interface, aprimorando a funcionalidade e principalmente a estética do projeto, resultando em uma experiência mais completa e amigável para os usuários. As imagens 10, 11 e 12 mostram respectivamente as telas de envio de arquivo, *dashboard* e configurações de referências para todos os exames.

Figura 10: Interface do usuário na tela inicial da aplicação, envio de arquivos.



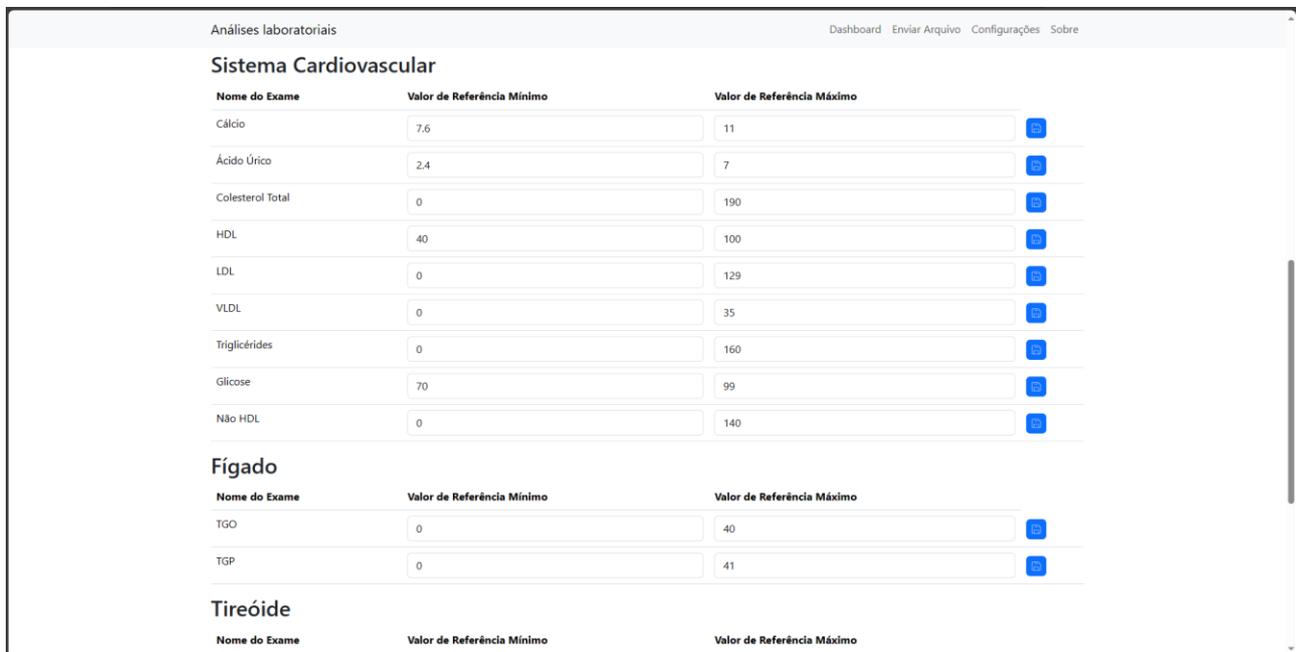
Fonte: Dos autores.

Figura 11: Interface do usuário na tela principal da aplicação, *dashboard*.



Fonte: Dos autores.

Figura 12: Interface do usuário na tela de configurações de referências.



Fonte: Dos autores.

Na aplicação desenvolvida, quando o usuário faz o upload de um documento, o processo é iniciado com uma requisição ao servidor. Nesse ponto, o sistema realiza as conversões necessárias para otimizar a

extração precisa do conteúdo do laudo. Essa etapa de preparo é fundamental para garantir que o Tesseract possa extrair o conteúdo com a máxima precisão possível.

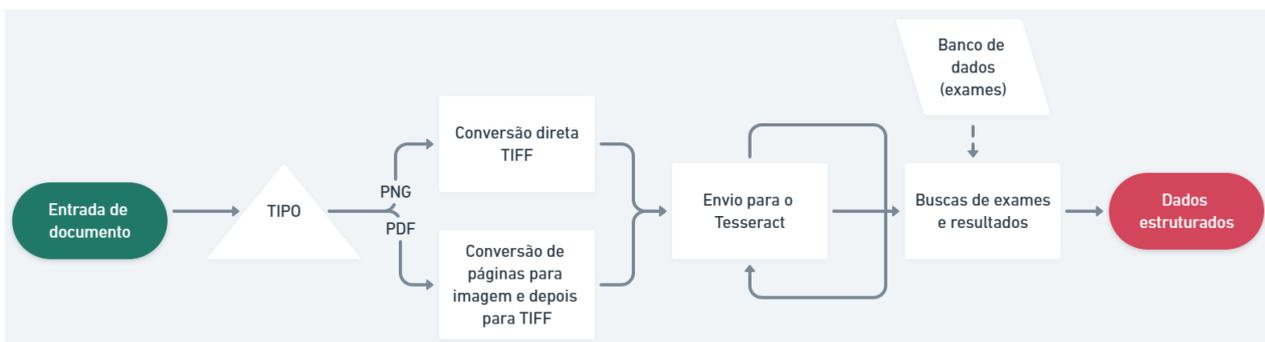
Em um nível mais baixo, o envio de um arquivo faz acesso a um endpoint do servidor web “/EnvioArquivo” que após receber esse documento faz a verificação de formato – se for uma imagem (PNG ou JPG, por exemplo) ela é convertida diretamente para TIFF, mas se for um PDF passa por uma conversão anterior, para que cada página desse PDF se torne uma única imagem – e como o módulo do Tesseract usado é em Python, foi necessário uma adaptação no servidor para que pudesse trabalhar com essa linguagem. As imagens são todas armazenadas numa pasta temporária após a conversão e só são deletadas no momento imediatamente anterior ao envio das informações de volta para o cliente.

A adaptação consiste na criação de um ambiente virtual para python (Venv, ou *virtual enviroment*), que é uma ferramenta nativa obtida junto às ferramentas da linguagem, para que pacotes possam ser instalados de forma local e não globalmente no sistema. Após os devidos ajustes mencionados para a imagem, o servidor executa um comando, como se fosse uma entrada no terminal, para ativar o ambiente e liberar o acesso à ferramenta de leitura. Esses comandos todos foram encapsulados num bash (*shell script*) para modularizar melhor esse processo.

Cada imagem, ou cada página é lida individualmente pelo OCR e os resultados vão sendo gravados dentro de uma lista com todos os caracteres, linha a linha, de forma que após o fim do ciclo se tenha em uma única variável todas as informações convertidas na mesma formatação (mas agora acessível, podendo se realizar operações com *strings* por exemplo). A figura 13 mostra o fluxo geral da entrada do documento pelo lado cliente da aplicação até a geração do objeto de saída com os dados estruturados, posterior ao fim do fluxograma ficam ainda a montagem e armazenamento dos dados no front-end.

As imagens de entrada são convertidas para um formato que seja ideal para o módulo, geralmente o formato TIFF, conforme especificado na documentação da ferramenta. O uso desse formato é uma prática recomendada para maximizar a qualidade da extração de texto, uma vez que tem suporte a alta resolução e compressão sem perda de qualidade. A escolha do formato adequado é essencial para que o Tesseract possa processar de forma eficaz o conteúdo dos exames laboratoriais, garantindo resultados confiáveis na etapa seguinte e contribuindo para a qualidade de toda a cadeia.

Figura 13: Fluxograma parcial: entrada do arquivo até o envio para o Tesseract.



Fonte: Dos autores.

Internamente no módulo do OCR são feitas algumas correções e tratamentos na imagem com a biblioteca OpenCV (em tradução livre, “Biblioteca aberta de visão computacional”), por exemplo: “`cv2.resize, fx = 2, fy = 2, interpolation = cv2.INTER_CUBIC`” que vai redimensionar a imagem para um tamanho maior, usando um fator de 2 para ambos os eixos e interpolação cúbica, que é uma técnica que oferece boa qualidade e é adequada para esse tipo de aumento.

O servidor aguarda até que todas as imagens (da pasta temporária) sejam lidas e armazenadas na lista correspondente, para que só então se inicie a busca pelos resultados dos exames. A primeira ação realizada é a tentativa de buscar alguma informação do laudo que remete aos laboratórios que possuem alguma regra específica criada para seu formato. Caso a busca seja realizada com sucesso, o algoritmo aponta para o conjunto

de regras de leituras específicas para os exames; caso falhe, aponta para um caso *default*, que também vai buscar as informações, porém com uma precisão menor.

De forma independente do resultado da busca pelo laboratório, os exames que foram definidos previamente no banco de dados, nesse momento são lidos para que possam ser buscados dentro da cadeia de caracteres criada. A busca acontece de forma linear, ao encontrar com o exame específico, exemplo: “Hemoglobina”, procura-se um valor numérico mais próximo à frente do título encontrado. Quando o valor é encontrado ele é adicionado em um objeto que será retornado ao fim da busca pelo último exame.

## Resultados

Os resultados obtidos pela ferramenta de interpretação de exames laboratoriais mostraram variações significativas, influenciadas por diferentes características das entradas submetidas à análise. Dentre os fatores determinantes, destacam-se a qualidade, o formato, a nitidez da imagem e a quantidade de informações fornecidas. Essas particularidades foram identificadas como pontos cruciais que influenciam a precisão e eficácia do processo de interpretação.

Uma das dificuldades centrais enfrentadas durante o desenvolvimento da aplicação foi a adaptação a diversos formatos e organização interna presentes nos documentos provenientes de diferentes laboratórios. Cada laboratório frequentemente apresentava uma estrutura única para seus relatórios de exames, o que demandou uma adaptação específica para cada um. Para enfrentar esse desafio, foram coletados e analisados exames provenientes de laboratórios reconhecidos, como a Unilab (laboratório da Unimed), para identificar padrões recorrentes e também laboratórios chamados de “genéricos”, aqueles que não foram adicionadas regras específicas. Com base nesses padrões identificados, regras específicas para após a leitura foram desenvolvidas e incorporadas à aplicação. Essas regras foram projetadas para permitir que a aplicação compreendesse a estrutura e os dados de cada tipo de exame, possibilitando uma interpretação mais precisa e adequada.

Além das regras específicas, foi implementado um processo de interpretação mais genérico. Esse método foi concebido para buscar os dados de maneira linear e direta, sem adaptações direcionadas para estruturas de laboratórios individuais. Essa abordagem foi implementada com o objetivo de garantir que, mesmo laboratórios com estruturas distintas e menos padronizadas pudessem ser interpretados pela aplicação, embora com uma precisão potencialmente menor.

Essa adaptação flexível e a inclusão de regras específicas, juntamente com um método genérico possibilitaram uma interpretação mais ampla e abrangente dos exames laboratoriais, buscando atender às diferentes necessidades e variações presentes nos documentos provenientes de diferentes fontes. Essas estratégias visam ampliar a utilidade da aplicação, permitindo que ela fosse aplicável a uma ampla gama de laboratórios, mesmo diante das diversidades estruturais e de informações apresentadas nos relatórios de exames.

Os problemas principais na extremidade final da aplicação são: não presença de exame no resultado e resultado incorreto. Na mostra disponível, os resultados da leitura de um PDF foram mostrados na tabela 1 abaixo, indicando se o exame foi identificado e correto (IC), identificado e incorreto (II), não identificado (NI) ou não disponível (ND), caso o documento não possua o exame correspondente.

Tabela 1: Resultados obtidos relacionando exame com diferentes laboratórios.

(continua)

	<b>Unilab</b>	<b>Sabin</b>	<b>São Lucas</b>	<b>Genérico 1</b>	<b>Genérico 2</b>
<b>Cálcio</b>	IC	ND	ND	NI	ND
<b>Hemoglobina</b>	IC	IC	IC	IC	II
<b>VCM</b>	IC	IC	IC	IC	IC
<b>HCM</b>	II	IC	II	IC	IC
<b>CHCM</b>	NI	IC	IC	IC	IC
<b>RDW</b>	IC	IC	IC	IC	II
<b>Leucócitos</b>	IC	IC	IC	II	II
<b>Eosinófilos</b>	IC	IC	IC	IC	NI
<b>Basófilos</b>	IC	IC	IC	IC	II
<b>Linfócitos</b>	II	IC	IC	II	II
<b>Linfócitos Atípicos</b>	IC	ND	ND	IC	IC
<b>Monócitos</b>	II	IC	IC	IC	II
<b>Plaquetas</b>	IC	IC	IC	IC	IC
<b>Hematócrito</b>	IC	IC	IC	IC	II
<b>Ácido Úrico</b>	IC	ND	IC	ND	ND
<b>Colesterol Total</b>	IC	ND	ND	NI	II
<b>HDL</b>	IC	ND	ND	II	NI
<b>LDL</b>	IC	ND	ND	ND	NI
<b>VLDL</b>	IC	ND	ND	ND	NI
<b>Triglicérides</b>	IC	ND	ND	II	II
<b>Creatinina</b>	IC	ND	IC	II	II
<b>Glicose</b>	II	ND	ND	ND	II
<b>TGO</b>	IC	ND	ND	II	ND
<b>TGP</b>	IC	ND	ND	II	ND
<b>T3 Livre</b>	ND	ND	ND	II	ND
<b>T4 Livre</b>	IC	ND	ND	II	II

Tabela 1: Resultados obtidos relacionando exame com diferentes laboratórios.

(continuação)

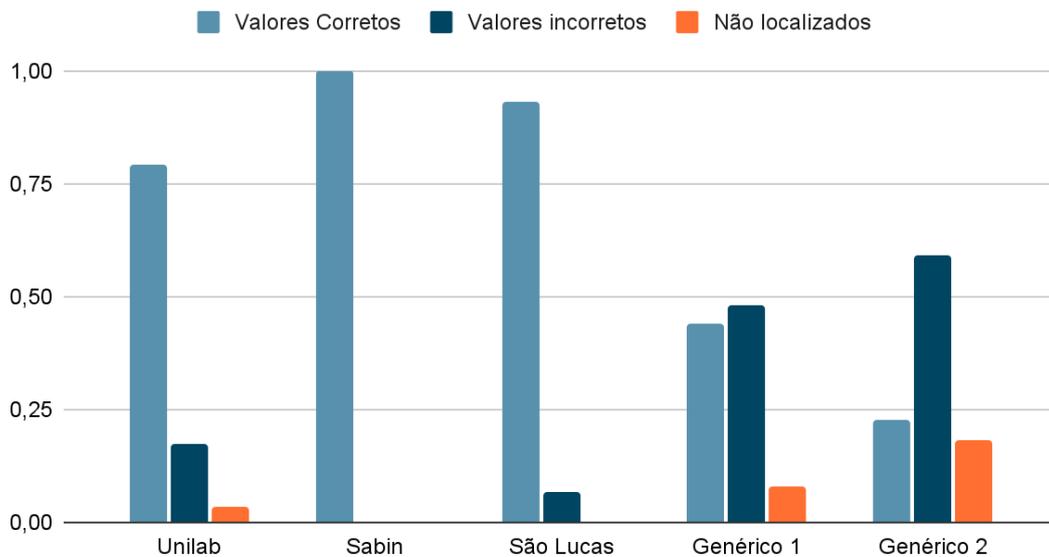
	Unilab	Sabin	São Lucas	Genérico 1	Genérico 2
<b>TSH</b>	IC	ND	ND	II	II
<b>Uréia</b>	IC	ND	IC	II	ND
<b>Não HDL</b>	IC	ND	ND	ND	ND
<b>Hemoglobina Glicada</b>	II	ND	ND	II	ND

Fonte: Dos autores.

Considerando somente os exames existentes (não “ND”) em cada arquivo submetido, as taxas de sucesso foram indicadas em porcentagem no gráfico abaixo (figura 14).

Figura 14: Taxa de acerto do algoritmo para diferentes laboratórios.

### Porcentagem de acerto por laboratório



Fonte: Dos autores.

## Discussão

Como esperado, os exames chamados “genéricos”, ou seja, que não possuem regras específicas para adaptação da estrutura possuem uma taxa de acerto menor para valores corretos quando comparados com os demais, foi pensando nisso que foi adicionada a funcionalidade de edição do valor pela interface gráfica no sistema.

De maneira geral, os problemas com valores incorretos e falha na localização podem ser explicados pela tradução incorreta de algum caractere, isso pode acontecer em qualquer caso, genérico ou específico devido às condições da entrada já mencionadas como qualidade, tamanho, ruído, etc.

Entre os exemplos apresentados, os problemas de não-localização de exames, ou seja, dentro da lista de caracteres convertidas não foram identificados os títulos dos exames, o único caso em que isso é possível é o caso de conversão incorreta de algum caractere, o que impede o algoritmo de realizar corretamente a comparação com a *string* vinda do banco de dados (essa causa é exclusiva).

Entre os valores incorretos (principalmente entre exames genéricos), nota-se que existe uma boa parcela de valores nulos (objeto com os valores de saída na figura 15), isso geralmente ocorre quando um título é localizado, mas não existe valor numérico nas proximidades (a causa não é exclusiva, pode ser falha na tradução também). Entretanto, observando o gráfico da figura 14 se pode concluir que as regras específicas para os laboratórios não genéricos funcionam bem. Ao que não se refere à tradução, indica uma formatação não linear dos resultados, ou seja, a linha que contém o título do exame não contém seu resultado e por isso o servidor foi incapaz de encontrar um resultado válido (numérico).

Figura 15: *Print* do terminal do objeto de saída com os resultados do processamento de um exame genérico.

```
Laboratório: Laboratório não encontrado

Caso default
{
  Leucocitos: null,
  Glicose: null,
  Calcio: null,
  Hemoglobina: 130,
  Hematocrito: 405,
  VCM: 96.4,
  HCM: 31,
  CHCM: 32.1,
  RDW: 132,
  Basofilos: 6,
  Linfocitos: 416,
  'Linfocitos Atipicos': 0,
  Monocitos: 78,
  Plaquetas: 187000,
  Creatinina: null,
  'Colesterol Total': null,
  Triglicerides: null,
  'T4 Livre': 4,
  TSH: null
}
```

Fonte: Dos autores.

A solução trivial para o problema de não encontro de valores na busca em linha, seria continuar a procura até encontrar um valor numérico (até a próxima linha por exemplo), todavia isso não pode ser implementado pois a maioria dos exames possuem valores de referência próximos a cada exame e como o algoritmo não possui esse tipo de distinção (somente é possível fazer verificação de tipos) acabaria sendo lido uma referência no lugar do resultado do exame.

## Considerações Finais

A implementação e desenvolvimento da ferramenta de interpretação de exames laboratoriais representa um avanço significativo no campo da saúde, fornecendo uma solução inovadora para profissionais da saúde lidarem com a análise e interpretação de exames bioquímicos. Durante o processo de criação desta aplicação, diversos desafios foram enfrentados, desde a variedade de formatos e estruturas dos documentos provenientes de diferentes laboratórios até a necessidade de adaptar a interpretação para cada tipo de exame.

A resposta a esses desafios foi a implementação de uma solução flexível e adaptável, capaz de lidar com diferentes formatos de entrada, seja por meio de regras específicas desenvolvidas para laboratórios conhecidos, ou por meio de um método mais genérico. Essas estratégias combinadas, buscaram oferecer um resultado mais preciso e abrangente dos exames, contribuindo para a prática clínica e melhorando a qualidade dos cuidados de saúde oferecidos.

É importante ressaltar que, embora a ferramenta tenha alcançado resultados significativos, há oportunidades contínuas de aprimoramento. A contínua evolução da aplicação, com a inclusão de novos algoritmos de processamento, a melhoria da interface do usuário e a incorporação de novos padrões de exames, pode aprimorar ainda mais a eficiência e precisão do sistema. Um detalhe importante do estado atual do modelo é que só são aceitos exames completos através de PDFs, já que atualmente só é possível submeter um único arquivo e geralmente não se tem uma imagem grande suficiente para conter todos os exames.

Além disso, a adaptação e expansão contínua da aplicação para lidar com novas fontes de exames e estruturas diferentes são áreas de foco para futuros aprimoramentos. Esse processo permitirá à aplicação ser mais abrangente e atender a uma gama ainda maior de necessidades clínicas, aprimorando assim a assistência à saúde.

Apesar da aplicação ter proporcionado de maneira eficiente a interpretação dos exames laboratoriais. A capacidade de segmentar e visualizar os resultados diferentes de diferentes grupos corporais em uma única interface facilitou a análise e posterior interpretação dos dados, não é possível, ainda, o acompanhamento histórico dos exames dos pacientes, a estrutura desenvolvida é capaz de suportar esse tipo de melhoria, fazendo com que o paciente consiga acompanhar sua evolução no tratamento ou recuperação de doenças que tenham relação com esses exames.

Em suma, a ferramenta desenvolvida demonstra o poder da integração entre tecnologia e medicina clínica, proporcionando um avanço significativo do campo. Com adaptações progressivas, a aplicação tem o potencial de desempenhar um papel fundamental na melhoria dos cuidados de saúde, oferecendo uma interpretação mais ágil, precisa e abrangente dos exames bioquímicos, beneficiando não apenas os profissionais de saúde, mas também os pacientes de forma geral.

## Referências

AMBRÓSIO, Paulo Eduardo. Redes neurais artificiais no apoio ao diagnóstico diferencial de lesões intersticiais pulmonares. 2002. Tese de Doutorado. Universidade de São Paulo. Disponível em: <https://www.teses.usp.br/teses/disponiveis/59/59135/tde-26102002-155559/publico/Dissertacao.pdf>. Acesso em: 10 de Outubro de 2023.

ANDRIOLO, A., et al. (2018). Recomendações da sociedade brasileira de patologia clínica/medicina laboratorial (SBPC/ML): fatores pré-analíticos e interferentes em ensaios laboratoriais. Editora Manole Ltda.

DE CASTRO, F. C. C.; DE CASTRO, M. C. F. Redes neurais artificiais. Porto Alegre, RS: Pontifícia Universidade Católica do Rio Grande do Sul, 2001.

FATEMAN, Richard J. et al. Optical character recognition and parsing of typeset mathematics 1. Journal of Visual Communication and Image Representation, v. 7, n. 1, p. 2-15, 1996.

GEEKSFORGEES. Introduction to Recurrent Neural Network. 18 de Maio de 2023. Disponível em: <https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/>. Acesso em: 20 de Out. de 2023.

HAYKIN, Simon. Redes neurais: princípios e prática. Bookman Editora, 2001.

ICMC-USP. Redes Multi Layer Perceptron treinadas com BackPropagation. Disponível em: <https://sites.icmc.usp.br/andre/research/neural/#cara>. Acesso em: 12 de Out. de 2023.

KOVÁCS, Zsolt László. Redes neurais artificiais. Editora Livraria da Física, 2002.

LIMA-OLIVEIRA G, Volanski W, Lippi G, Picheth G, Guidi GC. Pre-analytical phase management: a review of the procedures from patient preparation to laboratory analysis. Scand J Clin Lab Invest. 2017 May;77(3):153-163. doi: 10.1080/00365513.2017.1295317. Epub 2017 Mar 7. PMID: 28266238.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, v. 5, p. 115-133, 1943.

META, Platform. Inc. Menlo Park, Califórnia, Estados Unidos, 2004.

MITHE, Ravina; INDALKAR, Supriya; DIVEKAR, Nilam. Optical character recognition. International journal of recent technology and engineering (IJRTE), v. 2, n. 1, p. 72-75, 2013.

NAGY, George. At the Frontiers of OCR. Proceedings of the IEEE, v. 10, n. 7, Julho de 1992. Disponível em: [https://sites.ecse.rpi.edu/~nagy/PDF\\_files/Nagy\\_ProcsIEEE92-Frontiers-Of-OCR.pdf](https://sites.ecse.rpi.edu/~nagy/PDF_files/Nagy_ProcsIEEE92-Frontiers-Of-OCR.pdf). Acesso em: 15 de Out. de 2023.

RAUBER, Thomas Walter. Redes neurais artificiais. Universidade Federal do Espírito Santo, v. 29, 2005. Disponível em: [https://www.researchgate.net/profile/Thomas-Rauber-2/publication/228686464\\_Redex\\_neurais\\_artificiais/links/02e7e521381602f2bd000000/Redes-neurais-artificiais.pdf](https://www.researchgate.net/profile/Thomas-Rauber-2/publication/228686464_Redex_neurais_artificiais/links/02e7e521381602f2bd000000/Redes-neurais-artificiais.pdf). Acesso em: 16 de Out. de 2023.

SINGH, Amarjot; BACCHUWAR, Ketan; BHASIN, Akshay. A survey of OCR applications. International Journal of Machine Learning and Computing, v. 2, n. 3, p. 314, 2012. Disponível em: <http://www.ijmlc.org/papers/137-L0022.pdf>. Acesso em: 18 de Out. de 2023.

SOUSA, A. C. N.; RODRIGUES JUNIOR, O. M. . Main errors in the pre-analytical phase of laboratory tests: an integrative literature review. Research, Society and Development, [S. l.], v. 10, n. 15, p. e261101523662, 2021. DOI: 10.33448/rsd-v10i15.23662. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/23662>. Acesso em: 20 de Out. de 2023.

STAUEMEYER, Ralf C.; MORRIS, Eric Rothstein. Understanding LSTM - a tutorial into Long Short-Term Memory Recurrent Neural Networks. 23 de Setembro de 2019. Disponível em: <https://arxiv.org/pdf/1909.09586.pdf>. Acesso em: 15 de Out. de 2023.